

---

# Recreating Human Mobility Patterns Through the Lens of Social Media: Using Twitter to Model the Social Ecology of Crime

James C. Wo<sup>1</sup> , Ethan M. Rogers<sup>1</sup>, Mark T. Berg<sup>1</sup>, and Caglar Koylu<sup>1</sup>

## Abstract

Studies of neighborhood crime are often limited in their ability to account for the dynamic nature of human mobility, a central tenet of prominent theoretical perspectives on the spatial distribution of crime. Yet, recent work indicates the utility of social media data for estimating the size and composition of such ambient population. In the present study, we assess whether four Twitter-derived measures are associated with crime counts across 2,348 block groups. Specifically, we focus on the density of Twitter users (and tweets), as well as the proportion of Twitter users (and tweets) that are “insiders.” We inferred Twitter users’ “insider” location from the block group in which they tweeted most frequently.

## Keywords

neighborhood crime, routine activities, social media, ambient population, guardianship

---

<sup>1</sup>The University of Iowa, Iowa City, USA

## Corresponding Author:

James C. Wo, University of Iowa, 400T North Hall, Iowa City, IA 52242-1223, USA.  
Email: james-wo@uiowa.edu

## Introduction

The assumption that crime is more frequent in more populated locations “is one of the few accepted ‘facts’ in criminology” (Boivin & Felson, 2018, p. 466). Of consequence, nearly all neighborhood studies of crime account for the residential population of an area, either by calculating crime rates as the outcome or by including population size as a statistical control (Chamlin & Cochran, 2004). The residential population, however, is unlikely to fully capture the social ecology of an area, as it pertains to crime patterns. This is largely attributed to human mobility: that is, people spend a significant amount of time outside of the spatial unit in which they reside—they travel to other spatial units for work, education, leisure, and consumption activities. Many neighborhoods are thus occupied by a *mélange* of residents and temporary users of the space whom collectively constitute the “ambient population” (Boivin & Felson, 2018, p. 469), all of whom are important to the production of local crime rates (e.g. Stults & Hasbrouck, 2015).

Traditionally, scholars of neighborhoods and crime have utilized administrative data from the Census Bureau and other agencies to measure characteristics of spatial areas in terms of their physical environment, economic well-being, and residential population (Hipp et al., 2019). These measures such as poverty, income, and residential mobility are then linked to aggregate crime outcomes under the assumption that they effectively capture criminal opportunity and social disorganization mechanisms (Pratt & Cullen, 2005; Sampson et al., 1997; Stucky & Ottensmann, 2009). Characteristics of the residential population and the physical environment, however, remain limited in their ability not only to account for the range of persons occupying an area for some period, whether residents or non-residents, but also to capture the degree of activity of both residents and non-residents within an area. These limitations are important because the density and activities of “outsiders”—non-resident infrequent visitors—might increase crime in places (P. L. Brantingham & Brantingham, 1995; Felson & Boba, 2010), whereas the density and activities of “insiders”—residents and frequent users of the space (e.g., workers)—may generate the opposite effect: reduced crime. We propose a strategy that distills the composition and activity of people within a spatial unit, and therefore might provide insights into how neighborhoods successfully (or unsuccessfully) minimize crime.

Prior studies have discussed the limitations of linking administrative measures to aggregate crime outcomes (Gerber, 2014; Hipp et al., 2019; X. Wang et al., 2012; M. Wang & Gerber, 2015); yet, accounting for the ambient population of spatial units has proved to be a bedeviling challenge in criminology and urban sociology. We focus on the utility of social media data to capture

conceptualizations of the ambient population. Recent research has demonstrated that social media holds promise as a tool to better understand the ambient population of communities beyond that of measures obtained from administrative data (Hipp et al., 2019; Malleon & Andresen, 2015a). For all their promise, measures derived from social media and other human activity platforms are measured globally and therefore typically do not account for the composition of the ambient population of spatial area. These measures thus include people who reside in the space or visit the space frequently (we refer to as “insiders”), and those who visit the space irregularly but are there occasionally for work or leisure (we refer to as “outsiders”). The distinction between the usage groups is important, however, because the people in each theoretically have overlapping interests in the local community and its well-being. To advance this emerging line of work it is therefore important to move beyond global measures to decompose the ambient population of social media indicators into the proportion of insiders and outsiders who occupy spatial units or neighborhoods.

The current study employs 9 months of geo-tagged Twitter data from a sample of users in Los Angeles to estimate the ambient population of the city’s Census block groups and the proportion of users who are insiders as opposed to outsiders. These measures are then used to predict crime in 2,348 Los Angeles block groups over the 9-month period. Thus, the study pursues two specific aims: *First*, we specify measures of local human activity as alternative estimates of the ambient population of Census block groups that distinguishes between occupants of a neighborhood who are likely insiders as opposed to outsiders. Note, however, that this distinction does not isolate residents from non-residents, but rather it assesses the extent to which a space is occupied by *frequent* users, who may or may not include resident inhabitants, rather than *infrequent* users (or outsiders). *Second*, drawing upon assumptions from social disorganization and routine activities theories, we examine if the different Twitter measures of the ambient population have unique associations with both property and violent crime.

## Background

### *Ambient Population and Neighborhood Crime*

Nearly every major criminological theory which accounts for the social ecology of places considers the residential composition of places as an important variable in the explanation of crime patterns. These theories generally conceive of the local population, however measured, as a metric for specifying local crime risk or exposure (see Osgood & Chambers, 2000). To estimate the

population at risk within a neighborhood, most studies use Census estimates of the number of individuals who reside in households within a given area (Boivin & Felson, 2018). This measurement strategy is often adopted for convenience and accessibility given that alternative methods of assessing the local ambient population are difficult to ascertain. Scholars who study the spatial distribution of crime, however, have long raised concerns about the utility of administrative-based measures (Andresen, 2006; Boggs, 1965; Cohen et al., 1985). As noted, the residential population does not reflect the dynamic nature of human mobility—that is, people spend a significant amount of time outside of the spatial unit (Hipp et al., 2019; Mburu & Helbich, 2016; Stults & Hasbrouck, 2015).

Although the dynamic population receives rather limited empirical attention, human mobility is central to routine activities theory (Cohen & Felson, 1979) and its geographic corollary, crime pattern theory (P. J. Brantingham & Brantingham, 1984; P. L. Brantingham & Brantingham, 1995). According to these perspectives, the amount of crime a place experiences is determined largely by the level and type of human activity that occurs there (Reynald, 2011). Administrative measures of the residential population are therefore limited in their ability to account for the full range of persons occupying an area for some period of time. As such, to understand the spatial distribution of crime, researchers benefit from knowing the “ambient population” of an area—the number of people in an area at any given time of the day or year, regardless of where they live (Andresen, 2006, p. 259; Boivin & Felson, 2018, p. 469).

As noted, researchers have increasingly used a novel collection of proxy or indirect measures to capture the ambient population beyond those traditionally derived from the Census. Such indirect estimates have included land use characteristics (Bernasco & Block, 2011; Boessen & Hipp, 2015; Haberman & Ratcliffe, 2015) and the number of employees in an area (Boivin, 2013; Hipp et al., 2017; Wo, 2016). Other researchers have adjusted Census estimates of the residential population based upon alternative measures of land use characteristics and activities (e.g., Andresen, 2006). For example, using satellite imagery from the LandScan Global Population Database, Andresen (2006) used an average 24-hour ambient population that adjusts the residential population according to the “relative attractiveness” of the location, such as road proximity, slope, land cover, and nighttime lights (see also Andresen, 2011; Andresen & Jenion, 2010). Other more direct estimates of the ambient population have been calculated using transportation or commuter data (Boivin & Felson, 2018; Felson & Boivin, 2015; Mburu & Helbich, 2016; Stults & Hasbrouck, 2015), mobile phone data (Hanaoka, 2018; Song et al., 2018), and social media data (Hipp et al., 2019; Malleson & Andresen 2015a, 2015b, 2016).

Across these studies, the conclusions are often very similar—“residential and ambient populations are non-redundant predictors of crime” (Boivin & Felson, 2018, p. 468). Stults and Hasbrouck (2015), for instance, find that city rankings of crime rate estimates vary considerably when accounting for the daily working population in the denominator of crime rates. Similarly, Andresen (2006, 2011) shows that crime rate maps in Vancouver differed when using the ambient population rather than the residential population. Social media data produce similar findings. Using crowd-sourced social media data as the indicator of the ambient population, Malleson and Andresen (2015a) reveal that some neighborhoods in Leeds, United Kingdom have a comparatively high violent crime rate when using the residential population denominator, but have an average violent crime rate after accounting for the ambient population. A recent study by Hipp et al. (2019) using geolocated Twitter data for approximately 100,000 blocks in Southern California revealed that the ambient population during 2-hour time periods was positively associated with the level of crime during the same 2-hour time period. This particular study is important because it demonstrates that the ambient population affects crime independent of the residential population (Hipp et al., 2019).

### *Insider and Outsider Ambient Population*

A persistent challenge to research on the ambient population is that estimates often consist of both residents and non-residents or insiders and outsiders (Boivin & Felson, 2018; Hipp et al., 2019). Prominent theoretical work indeed suggests the composition of the ambient population may be important for explaining the spatial distribution of crime. As already noted, routine activities perspectives emphasize the role of human mobility and offer a dynamic approach to explaining the spatial distribution of crime—the spatial and temporal patterns of daily activities involving family, work, or leisure explain the distribution of criminal opportunity and, in turn, crime (Cohen & Felson, 1979). Criminal opportunities arise where motivated offenders converge with suitable targets in the absence of capable guardianship (Cohen & Felson, 1979; Felson, 1987). In other words, more offenders and targets in an area should correlate with more crime, while more guardians in an area should correlate with less crime. The challenge when considering the ambient population, however, is that the composition of the ambient population is often unclear. Indeed, the ambient population “has a complicated relationship with the presence of offenders, targets, and guardians. . .” (Hipp & Williams, 2020, p. 81).

Upon closer inspection, the literature suggests that the conceptual distinction between members of the ambient population in terms of their residency status may miss the importance of *how* people use a space or a neighborhood.

A person who spends a significant amount of their time outside of their own residential neighborhood and inside another neighborhood to work, attend school, or visit family and friends would, in theory, have a strong investment in the well-being of the community similar to the sentiment of a local resident. Jacobs (1961) “new urbanism” vision of the local community positioned crime control efforts as an investment shared by permanent residents and frequent visitors who value the space for its material and social resources and the functional coordination of human and business interests. The same sentiments of attachment and investment may not be carried by actors that use the spatial area less frequently, including infrequent patrons to local businesses, occasional visitors, and those passing through temporarily from one location to another. Not only may infrequent users be weakly tethered to parochial and primary social networks but they may also maintain a degree of social anonymity by virtue of their intermittent presence (Bursik, 1999; Hunter, 1985). To the extent that these features reduce the barriers to criminal behavior, then an ambient population comprised of a higher proportion of infrequent users might increase the amount of crime in an area. For these reasons, perhaps a more accurate conceptualization of the ambient population would divide the occupants of a space into insiders and outsiders, with the former category containing both residents and non-residents who frequently inhabit a neighborhood for work, leisure, or other human activities (Hunter, 1985). Here forward this conceptualization guides the current framework.

There are two interpretations of the way the composition of the ambient population might relate to an area’s risk for crime—each offers a different account of the contributions of insiders and outsiders to the degree of risk. On the one hand, the composition of the ambient population may not matter for understanding the spatial distribution of crime. Insiders as well as outsiders, just as residents as well as non-residents, are potential offenders, potential targets, and potential guardians all at the same time (Boivin & Felson, 2018). For example, both insiders and outsiders act as guardians by their mere presence (Felson & Boba, 2010; Hollis-Peel et al., 2011). Through their simple proximity and visibility, or “natural surveillance,” even strangers can discourage crime (Newman, 1972). By being “on the scene,” any bystander could deter a motivated offender through signaling the *possibility* that a crime event would be detected and an intervention would occur (Hollis-Peel et al., 2011, pp. 56, 57). Accordingly, “it is the idea that someone is watching and *could* detect problematic behaviors or people that deters the likely offender from committing a criminal act” (Hollis-Peel et al., 2011, p. 66). In a similar view, as noted above, the “New Urbanism” literature suggests that non-residents, just like residents, provide consistent “eyes on the street” and, in turn, represent the potential for increased guardianship rather than increased crime

(Achimore, 1993; Hillier, 2004; Jacobs, 1961). Thus, the *size* of the ambient population, regardless of insider status or residency, may be associated with both increases (due to increased offenders and targets) and decreases (due to increased guardians) in the number of crimes.

On the other hand, more recent work emphasizes that insiders—residents and frequent visitors—may be more effective guardians than non-resident outsiders and that outsiders may even interfere with the effectiveness of guardianship within a spatial unit. Insiders may be more effective guardians than outsiders because they are more likely to feel responsible for, and invested in, a specific neighborhood. When potential guardians feel greater responsibility for the people and places within a neighborhood, they will be more effective at deterring criminal events (Clarke, 1992; Felson, 1995, p. 57). In an extension of the guardianship concept of routine activities theory, Felson (1995) describes four levels of responsibility: personal, assigned, diffuse, and general (see also Clarke, 1992). Personal responsibility for crime discouragement is taken by those who own potential targets (e.g., homeowners) or are intimately related to owners (e.g., family, friends, and neighbors). Assigned and diffuse responsibility refers to the responsibility of employees to discourage crime either through specific assignment or by unintended consequences of specific job duties. Finally, general responsibility for crime discouragement is taken by “any bystander or visitor whose presence discourages crime or who notes an illegal activity that is or might be occurring there” (Felson, 1995, p. 56). According to Felson (1995), the tendency to discourage crime, as well as the directness and speed of discouragement, will be highest when guardians feel a personal or professional responsibility and will be lowest when guardians feel a general responsibility. Newman’s (1972) defensible space theory similarly emphasizes the association between crime at a location and “the number of users and extent of their felt responsibility” (p. 109).

In addition to greater feelings of responsibility, insiders may also be more effective guardians than outsiders because they are more capable of surveilling, monitoring, and supervising a spatial area (Reynald & Elffers, 2009). Indeed, the bulk of social disorganization literature focuses on the networks, informal social control efforts, and collective efficacy among the residents of a neighborhood (Bursik & Grasmick, 1993; Sampson et al., 1997; Sampson & Groves, 1989; Wilcox et al., 2004). People who are frequent visitors, however, may also form meaningful networks that can serve as the basis for effective guardianship within a neighborhood (Jacobs, 1961; Talen, 2002). Insiders who may not reside in a neighborhood, such as store owners, workers, and students, may still have the familiarity that facilitates surveillance and control efforts. Outsiders, in contrast, are unlikely to engage in extensive social

control efforts (Greenberg et al., 1982; Taylor et al., 1995; Wilcox et al., 2004) and are likely to be less sensitive to social controls (Boivin & Felson, 2018).

Theories of crime and space also suggest the presence of outsiders or infrequent users of the space may also inhibit the guardianship behaviors by insiders (Newman, 1972; Reynald & Elffers, 2009; Roncek, 1981; Taylor et al., 1995). The presence of outsiders lessens familiarity and makes it more difficult for insiders to identify potential offenders and detect suspicious activity (Reynald, 2011; Reynald & Elffers, 2009; Roncek, 1981; Taylor et al., 1995). In addition to the “blanket of anonymity” created by outsiders (Zahnow, 2018, p. 1122), the presence of non-residents may also discourage “territorial behavior in the form of guardianship” (Reynald, 2011, pp. 118, 119). Both residents and frequent visitors may feel less attached to and less responsible for their neighborhoods if non-resident outsiders contribute to physical and social deterioration (McCord et al., 2007). This withdrawal of insiders may, in turn, reduce the willingness and capacity of others to engage in guardianship behavior, such as surveillance and monitoring (Browning et al., 2017; Reynald, 2011; Reynald & Elffers, 2009; Taylor et al., 1995). Outsiders may also create “holes” in the insider-based fabric “for which no resident will take responsibility” (Taylor et al., 1995, p. 122; Wilcox et al., 2004).

Emerging evidence indicates that the ratio of insiders to outsiders in the ambient population may be associated with crime in neighborhoods. Land-use studies, for instance, tend to suggest that greater levels of crime are associated with land use patterns that presumably increase the ratio of non-residents to residents, such as non-residential land use (Bernasco & Block, 2011; Boessen & Hipp, 2015; Stucky & Ottensmann, 2009; Wo, 2019b) and mixed land use (Browning et al., 2010; Wo, 2019a; Zahnow, 2018). Mixed land use, for example, attracts outsiders and insiders to use its educational, commercial, and recreational services (P. L. Brantingham & Brantingham, 1995; Felson & Boba, 2010), and, in turn, presumably weakens insider-based control and undermines the ability to detect suspicious behavior (Taylor et al., 1995; Wo, 2019a; Zahnow, 2018). Indeed, Wo (2019a) found that neighborhoods with more land use heterogeneity tended to experience higher robbery and burglary rates, particularly among socioeconomically advantaged neighborhoods. Land use measures, however, only indirectly capture the composition of the ambient population.

Recent studies have assessed the presence of non-resident or outsider activities compared to insider activities based on data obtained in travel surveys. This research also lends support to the assumption that the ratio of insiders to outsiders should be negatively associated with crime (Boivin & Felson, 2018; Browning et al., 2017; Felson & Boivin, 2015). For instance,



using a sample of 192 census tracts, Browning et al. (2017) estimate the prevalence of non-residents using microsimulations of travel patterns. Results suggest that heightened rates of property crime are associated with a higher prevalence of non-residents. Using transportation data from 506 census tracts in a large Canadian city, Felson and Boivin (2015) examined the association between non-residents and the number of violent and property crimes. The results show that daily visits, particularly recreational visits, are associated with heightened numbers of both violent and property crimes. Using the same data, Boivin and Felson (2018) find that the inflow of non-residents is associated with both resident and non-resident crime.

## **Current Study**

The current study attempts to further clarify the extent to which both the size and composition of the ambient population is related to crime in geographic areas. Specifically, using novel social media data obtained from Twitter to assess human activity in Census block groups, we examine how human mobility patterns affect the spatial distribution of violent and property crimes. Further, drawing from our interpretation of routine activities and social disorganization perspectives and related theoretical work, we test whether the composition of the ambient population—defined here as insiders and outsiders—in Census block groups is associated with its counts of property and violent crime. Specifically, we hypothesize that the greater presence of insiders in the ambient population will be associated with lower levels of crime within a block group.

## **Methods**

### *Study Area and Units of Analysis*

The city of Los Angeles serves as the research setting for several reasons. First, routine activities and social disorganization theories generally view neighborhood effects on crime operating within an urban ecological system. Second, among prior studies to link ambient population or land use characteristics to crime across spatial units, most have been conducted in urban areas (Andresen, 2011; Browning et al., 2017; Malleson & Andresen, 2016; Wo, 2019b; Zahnaw, 2018). Finally, in order to produce robust measures of twitter activity and presence it is necessary to draw on a large sample of spatial units from a populous city (or area).

The block group is the second smallest geographic unit by which the U.S. Census Bureau aggregates data from surveyed households. An advantage of

using block groups as units of analysis is that the U.S. Census Bureau has created block groups to be internally homogenous on a range of sociodemographic characteristics, including income, race/ethnicity, educational attainment, and family structure. Thus, the block group has been widely used to examine neighborhood effects on crime. For the present study, we employ block groups as the units of analysis, and hereafter we refer to block groups and neighborhoods interchangeably.

### *Crime Data and Measures*

We collected data on crime incidents from the Los Angeles Police Department (LAPD) for the year 2014. The LAPD provided information on the location of such incidents as well as the type of crime that was committed. We geocoded these incidents using a geographic information system and aggregated them to block groups—the match rate was about 98%. To assess differences in crime across block groups, we have created the following outcome measures: a *violent crime index* (count of homicides, robberies, and aggravated assaults) and a *property crime index* (count of burglaries, larcenies, and motor thefts). Table 1 presents the summary statistics of all measures used in the analyses.

### *Twitter Data Collection and Processing*

Using the Twitter Streaming API, we collected geo-located tweets in the bounding box of Los Angeles County, California during a 9-month period between March 1, 2014 and December 1, 2014. Metadata of each tweet is also provided by the Streaming API, which include, but are not limited to, the textual content and the links for multimedia content of the tweet, the latitude and longitude coordinates, place-tags (if it is available), the date and time of the tweet and the local time zone, the user identification and name, the distinct tweet identification number, and the source application that is used to produce the tweet (e.g., Apple iPhone, Android, weather, news, government, agencies, and other non-personal Twitter user accounts).

Geo-located tweets are identified on two levels of geographic scale: (1) exact geographic coordinates with latitude and longitude coordinates derived from GPS sensors (referred to as “geo-tagged”) or (2) a descriptive format which includes a listed place name such as a point of interest (POI), neighborhood, or city (referred to as “place-tagged”). Place-tagged tweets often provide an area such as the city name rather than a specific coordinate pair for a location and are therefore more difficult to identify within a Census block group. Thus, we selected our 9-month collection period based on the highest

**Table 1.** Descriptive Results.

Variables	Sample statistics	
	M	SD
Crime counts		
Violent crimes	5.99	7.09
Property crimes	33.18	29.33
Block group predictors		
Tweet density <sup>a</sup>	1.77	2.24
User density	0.14	0.21
Insider activity	58.87	12.42
Insider presence	10.42	5.28
Percent residential	68.49	26.69
Percent retail	6.49	8.69
Concentrated disadvantage	0.15	1.07
Residential stability	-0.27	0.95
Ethnic heterogeneity	45.78	17.74
Percent aged 15 to 29	22.58	9.03
Population density <sup>b</sup>	173.75	139.68
Spatially lagged predictors		
Tweet density	1.69	1.32
User density	0.14	0.17
Insider activity	58.61	7.23
Insider presence	10.02	2.60
Percent residential	68.45	15.76
Percent retail	6.38	4.55
Concentrated disadvantage	0.14	0.93
Residential stability	-0.26	0.71
Ethnic heterogeneity	46.17	13.79
Percent aged 15 to 29	22.52	5.83
Population density	163.88	102.40

Note. N (block groups) = 2,348. SD = standard deviation.

<sup>a</sup>Twitter density measures refer to per 100 m<sup>2</sup>.

<sup>b</sup>Population density is measured in hundreds per square mile.

volume of geo-tagged tweets after analyzing multiple years of Twitter data from 2013 to 2018. In this period, 93% of tweets were geo-tagged while only 7% were place-tagged. Note that the percentage of geo-located tweets that are geo-tagged decreased considerably after 2014 because Twitter reversed its policy in April 2015 and recorded GPS coordinates only for those users

who explicitly authorized the publication of their precise location (Leetaru, 2019). For example, between August 1, 2015 and August 1, 2016, only 14% of tweets were geo-tagged, while 86% of tweets were place-tagged.

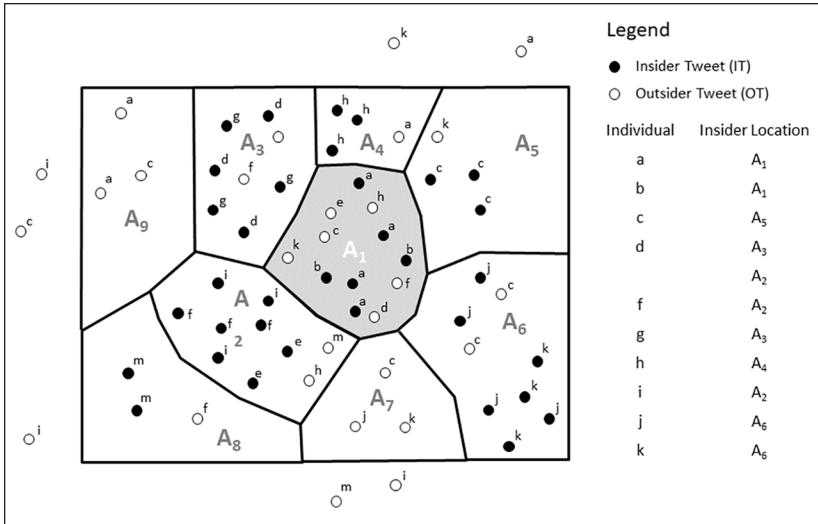
Using the metadata provided by the Twitter Streaming API, we excluded place-tagged tweets (4.28%), and the tweets authored by non-personal user accounts such as news feeds, weather and emergency reports, and external applications such as FourSquare and Instagram (9.39%). After the initial filtering, there were a total of 331,380 users who tweeted at least once within the border of Los Angeles City during the 9-month study period. These 331,380 users produced 12,039,954 geo-tagged tweets during the study period. About 75.90% of users (251,493) visited locations outside LA City and produced almost twice as many tweets outside LA City. Finally, we excluded any Twitter user who tweeted less than three times—about 90.04% (298,373) of users produced at least three tweets in LA City.

At the outset it is important to highlight some known limitations of using the available Twitter data, this includes (1) the sociodemographic representativeness of Twitter users (e.g., age, education, race, and ethnicity) relative to the general population (Nielsen, 2006; Tsou, 2015; Tufekci, 2014), (2) the rural urban gradient in Twitter user prevalence (Quattrone et al., 2015), (3) the user contribution bias caused by a few highly active users generating the majority of tweets (Nielsen, 2006), and (4) the representativeness of users who share their locations to all Twitter users (Pavalanathan & Eisenstein, 2015). We attempt to minimize the possible effects of these limitations by relying on social media data from a large diverse United States city with a rather large number of Twitter users during a particularly active period of Twitter usage.

### *Insider and Outsider Classification*

When a user's tweets are compiled over an extended period of time, it is possible to derive mobility and activity patterns of the user and thus, the ambient population for each neighborhood. We categorize the ambient population into insiders and outsiders, with the former category conceptualized as containing both residents and non-residents who frequently inhabit a neighborhood for work, leisure, or other human activities. Based on the sociological notions of places and place attachment (Oldenburg, 1999), we defined a user's insider neighborhood as the specific block group where the user spends most of their time (i.e., produces the majority of their tweets).

Figure 1 illustrates our strategy for determining a user's insider location. In this hypothetical schema, points represent locations of tweets, boundaries represent neighborhoods (block groups), and labels of points represent people

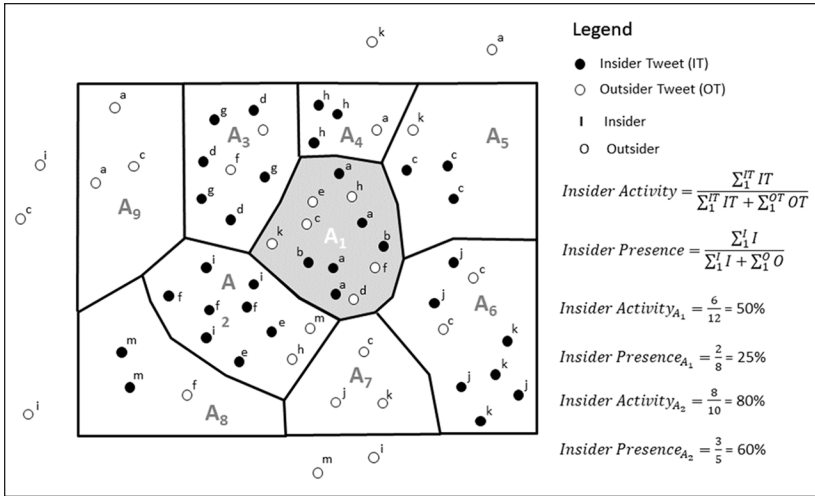


**Figure 1.** Insider and outsider tweet and user classification.

(e.g., *a*, *b*, and *c*). A tweet is classified as an “insider tweet” and symbolized with black if the tweet is produced by a user within their insider block group (i.e., the area in which the user produced the majority of their tweets). If a tweet is produced at a location other than the user’s insider block group, then the tweet is considered an “outsider tweet” and symbolized with white. For example, the most frequent tweet location for user *a* is area *A*<sub>1</sub>. The tweets that user *a* produced in *A*<sub>1</sub> are represented by black, whereas tweets that user *a* produced elsewhere (e.g., *A*<sub>4</sub>) are represented by white. Each user’s insider location is provided in the legend of Figure 1.

We then develop a set of metrics that approximate the relative density, activity and presence of insiders and outsiders within neighborhoods (block groups). The first two simple metrics are (1) *tweet density* and (2) *user density*, which are derived by dividing the total number of tweets and distinct users, respectively, by the area of each block group. The density metrics consider all users without distinguishing insiders and outsiders.

Figure 2 illustrates our strategy and formulas for computing insider activity and presence metrics using the same example represented in Figure 1. Our first metric is *insider activity*, which measures the proportion of all tweets that are produced by insiders. This metric is influenced by user contribution bias, as a large portion of tweets are produced by only a few users. Our second metric is *insider presence*, which measures the proportion of all users in



**Figure 2.** Metrics to measure insider activity and insider presence.

an area that are insiders. Insider presence considers each user in each area (block group) just once and addresses the user contribution bias. To illustrate, the insider activity of area  $A_1$  is calculated by dividing the number of tweets produced by insiders (i.e., six black tweets) by the total number of tweets within that area (i.e., six black tweets and six white tweets). In contrast, insider presence of area  $A_1$  is calculated by dividing the number of distinct insiders (i.e., two insiders, namely  $a$  and  $b$ ) by the total number of users within that area (i.e., eight users, namely  $a, b, c, d, e, f, h,$  and  $k$ ).

### *Neighborhood Socioeconomic Disadvantage, Mobility, and Population Characteristics*

To account for key socio-structural constructs from social disorganization theory (e.g., Sampson & Groves, 1989), we draw on American Community Survey 5-year estimates (2011–2015) of LA census block groups (such measures approximate the year 2013). We created a *concentrated disadvantage* scale based on a principal component factor analysis of four variables: (1) the percentage at or below 125% of the poverty level, (2) the percentage of single-parent households, (3) the average household income (reverse coded), and (4) the percentage with at least a bachelor's degree (reverse coded). This scale captures the degree of economic and social hardship that is linked to weakened informal controls (Sampson et al., 1997). We also created a

measure of *residential stability* by computing the mean of the standardized values of average length of residence (measured in years) and the percentage living in the same house 5 years ago. Prior studies have shown that residential instability impedes the development of mutual cohesion and trust among residents (Kubrin & Weitzer, 2003; Sampson & Groves, 1989; Shaw & McKay, 1942). To account for the level of *racial and ethnic heterogeneity* in neighborhoods, a Herfindahl index was created according to five racial and ethnic groups: white, Black, Latino, Asian, and other race (Gibbs & Martin, 1962). Racial and ethnic heterogeneity has been theorized to disrupt the dissemination of information and induce other components of social disorganization (Sampson & Groves, 1989). Furthermore, we constructed a measure of the residential population between 15 and 29 years old (i.e. *percent aged 15–29*) given that this reflects a crime-prone age group. Finally, we calculated *population density* (measured in hundreds per square mile) in order to account for the possibility that overcrowding might induce crime problems in neighborhoods (Hipp & Roussell, 2013; Shaw & McKay, 1942).

### *Land Use Characteristics*

Routine activities research has shown that certain land uses provide an abundance of criminal opportunities whereas others have the opposite effect. In particular, studies have largely demonstrated that commercial/retail land use heightens criminal opportunities whereas residential land use facilitates effective guardianship (Boessen & Hipp, 2015; Browning et al., 2010; Stucky & Ottensmann, 2009; Wo, 2019a). To account for the spatial distribution of criminal opportunities, we draw on 2012 administrative data from the Southern California Association of Governments to compute the percentage of the block group area classified into two land use measures: (1) *percent retail land use* (e.g. stores, restaurants, and shopping centers) and (2) *percent residential land use* (e.g., single family and multi-family).

### *Spatial Dependence*

Although social ecological factors of a neighborhood naturally impact crime in the focal neighborhood, prior studies have revealed the importance of accounting for the *broader spatial impact* of such factors (e.g., see Bernasco & Block, 2011; Boessen & Hipp, 2015; Peterson & Krivo, 2010). In other words, social ecological factors surrounding the focal neighborhood might exert substantive effects on crime in that neighborhood. We therefore constructed a spatially lagged measure for each of the independent variables using the first-order queen contiguity. For the focal block group, the spatially

lagged measure of concentrated disadvantage, for example, indicates the average level of disadvantage among contiguous block groups.

### Analytic Strategy

Negative binomial regression—an extension of Poisson regression—effectively accounts for an over-dispersed outcome variable via its dispersion parameter alpha (Hilbe, 2007). Given that the outcome variables of crime counts exhibit overdispersion, we employ negative binomial regression to examine neighborhood effects on crime across our sample of block groups with a nonzero population ( $N=2,348$ ).<sup>1</sup> A general expression of our estimated models can be expressed as follows:

$$y_{2014} = B_1 \mathbf{T}_{2014} + B_2 \mathbf{WT}_{2014} + B_3 \mathbf{S}_{2012} + B_4 \mathbf{WS}_{2012} + \alpha,$$

where  $y$  is the number of crimes in 2014,  $\mathbf{T}$  is a matrix of the twitter measures in 2014,  $\mathbf{WT}$  is a matrix of the spatially lagged twitter measures in 2014,  $\mathbf{S}$  is a matrix of the sociodemographic characteristic and land use measures in 2012,  $\mathbf{WS}$  is a matrix of the spatially lagged sociodemographic characteristic and land use measures in 2012, and  $\alpha$  is an intercept.

While one approach for modeling crime across spatial units is to specify population count as an exposure term (thereby estimating the outcome as a crime rate), we instead model crime counts by accounting for population density as a predictor, given concerns over population count being the denominator of a calculated crime rate (e.g., see Andresen & Jenion, 2010). We assessed and found no evidence of collinearity problems; for example, the variation inflation factor (VIF) scores did not exceed seven across all models.

The result section presents three models for both violent and property crime counts: (1) the effects of sociodemographic and land use characteristics without the Twitter measures, (2) the effects of *tweet density* and *user density*, and (3) the effects of *insider activity* and *insider presence*. These last two models control for sociodemographic and land use characteristics.

## Results

### Twitter Summary Statistics

A consideration of the descriptive statistics reveals several important patterns of Twitter use by individuals and block groups (not shown in tabular form). We find that individuals have a tendency to tweet from different block groups located across Los Angeles City. For instance, individuals tweeted from a



median of five distinct block groups over the 9-month period. The median number of insider tweets per user and the median number of outsider tweets per user was seven. Taken together, these summary statistics suggest that tweet activity is not infrequent nor spatially restricted (to home or non-home locations) with respect to the typical Twitter user in our dataset.

The descriptive statistics also indicate the median number of tweets made by insiders and outsiders from block groups was 1,778 and 1,134, respectively. These two medians suggest there is a considerable amount of Twitter activity across the LA block groups and therefore gives credence to using the Twitter platform to construct our activity measure—the percentage of all tweets posted by insiders. We also determine that rather large populations of users have typically been present in these block groups: the median number of insiders and outsiders were 21 and 197, respectively. These medians suggest that there is sufficient statistical power to distinguish between insider and outsider compositions across block groups over the 9-month period.<sup>2</sup>

### *Multivariate Regression Results*

Table 2 reveals the effects of the Twitter-derived measures and Census-based administrative measures on neighborhood crime. Models 1 and 4 of Table 2 provide the effects of the land-use and sociodemographic characteristics without the Twitter-derived measures. The land use measures, often included in models to approximate criminal opportunity and social (dis)organization mechanisms, are generally not related to the outcomes. This is inconsistent with prior studies which have largely found that residential and retail land uses are associated with various crime types (e.g., Boessen & Hipp, 2015; Wo, 2019b). These results underline, however, that our Twitter-derived measures may be useful proxies for the ambient population that appear to capture the dynamic nature of social ecology. Many of the socio-structural characteristics demonstrate significant effects on violent and property crimes in the expected directions. For instance, a one standard deviation increase in a focal block group's disadvantage is associated with a 21% increase in the expected number of violent crimes (model 1:  $b = .180$ ;  $SE = .04$ ;  $p < .01$ ) using the formula  $(\exp(\beta \times SD) - 1)$ .

In models 2 and 5 of Table 2, we find that measures of *tweet density* (block group and spatial lag) are predominantly not associated with violent crimes or property crimes. *User density*, however, tells a very different story. In terms of the focal block group, user density is related to more violent crimes (model 2:  $b = 0.857$ ;  $SE = 0.16$ ;  $p < .01$ ) and property crimes (model 5:  $b = 0.888$ ;  $SE = 0.12$ ;  $p < .01$ ) with all else being equal. When including the insider activity and insider presence measures, these effects not only remain

**Table 2.** Negative Binomial Regression Models of Crime Counts.

	Violent crime counts						Property crime counts					
	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
<b>Block group predictors</b>												
Tweet density	—		-0.015	(0.010)	-0.005	(0.010)	—		-0.014*	(0.007)	-0.011	(0.008)
User density	—		0.857**	(0.159)	0.709**	(0.157)	—		0.888**	(0.117)	0.821**	(0.117)
Insider activity	—		—		0.002	(0.002)	—		—		0.001	(0.001)
Insider presence	—		—		-0.019**	(0.004)	—		—		-0.006*	(0.003)
Percent residential	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)	0.000	(0.001)	-0.000	(0.001)	-0.000	(0.001)
Percent retail	-0.003	(0.002)	-0.003	(0.002)	-0.003	(0.002)	0.000	(0.002)	-0.000	(0.002)	-0.000	(0.002)
Concentrated disadvantage	0.180**	(0.039)	0.220**	(0.039)	0.219**	(0.039)	-0.051	(0.028)	-0.027	(0.028)	-0.028	(0.028)
Residential stability	-0.298**	(0.035)	-0.248**	(0.034)	-0.237**	(0.034)	-0.267**	(0.024)	-0.232**	(0.023)	-0.230**	(0.023)
Ethnic heterogeneity	0.003	(0.002)	0.003	(0.002)	0.003*	(0.002)	0.003**	(0.001)	0.003**	(0.001)	0.003**	(0.001)
Percent aged 15 to 29	0.005	(0.003)	0.002	(0.003)	0.003	(0.003)	0.005**	(0.002)	0.003	(0.002)	0.004*	(0.002)
Population density	-0.002**	(0.000)	-0.002**	(0.000)	-0.001**	(0.000)	-0.002**	(0.000)	-0.001**	(0.000)	-0.001**	(0.000)
<b>Spatially lagged predictors</b>												
Tweet density	—		-0.007	(0.022)	-0.003	(0.023)	—		-0.011	(0.017)	-0.002	(0.017)
User density	—		0.612**	(0.214)	0.630**	(0.224)	—		-0.172	(0.154)	-0.230	(0.161)
Insider activity	—		—		0.001	(0.004)	—		—		0.002	(0.003)
Insider presence	—		—		-0.005	(0.008)	—		—		-0.019**	(0.006)
Percent residential	-0.001	(0.001)	-0.002	(0.001)	-0.002	(0.001)	-0.003**	(0.001)	-0.003**	(0.001)	-0.003**	(0.001)
Percent retail	-0.006	(0.005)	-0.002	(0.004)	-0.002	(0.004)	-0.002	(0.003)	0.001	(0.003)	0.000	(0.003)
Concentrated disadvantage	0.577**	(0.048)	0.699**	(0.049)	0.684**	(0.050)	0.076*	(0.034)	0.121**	(0.034)	0.104**	(0.036)
Residential stability	-0.025	(0.051)	0.178**	(0.054)	0.168**	(0.056)	-0.091**	(0.035)	-0.013	(0.037)	-0.015	(0.038)
Ethnic heterogeneity	-0.001	(0.002)	0.000	(0.002)	-0.000	(0.002)	-0.002	(0.001)	-0.002	(0.001)	-0.002	(0.001)
Percent aged 15 to 29	-0.013**	(0.005)	-0.025**	(0.005)	-0.024**	(0.005)	-0.006*	(0.003)	-0.011**	(0.003)	-0.009**	(0.003)
Population density	0.000**	(0.000)	0.001**	(0.000)	0.001*	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Intercept	1.758**	(0.152)	1.937**	(0.147)	1.936**	(0.290)	3.828**	(0.104)	3.918**	(0.101)	3.948**	(0.201)

Note. N (block groups) = 2,348. b = unstandardized coefficient; SE = standard error.

\*p < .05. \*\*p < .01.

statistically significant, but they also are relatively strong in magnitude. There is a 16% increase in the expected number of violent crimes for a one standard deviation increase in the density of total users (model 3:  $b = 0.709$ ;  $SE = 0.16$ ;  $p < .01$ ). The equivalent comparison yields about a 19% increase in the expected number of property crimes (model 6:  $b = 0.821$ ;  $SE = 0.12$ ;  $p < .01$ ). We also observe a broader spatial impact of the density of total users on violent crimes. That is, a higher density of users in contiguous block groups is associated with more violent crimes in the focal block group—specifically, a 11% increase in violent crimes for a one standard deviation increase in such density (model 3:  $b = 0.630$ ;  $SE = 0.22$ ;  $p < .01$ ). There is no evidence of a broader spatial impact on property crime.

In models 3 and 6 of Table 2, we add the Twitter-derived measures of insider activity and insider presence to assess whether the composition of the ambient population is associated with neighborhood crime. The results suggest that insider activity (proportion of tweets made by insiders) in the focal block group and in contiguous block groups (spatial lag) is not associated with violent or property crime. Insider presence, however, is negatively associated with both outcomes. One standard deviation increase in the proportion of insiders is associated with 10% fewer violent crimes (model 3:  $b = -0.019$ ;  $SE = 0.00$ ;  $p < .01$ ) and 3% fewer property crimes (model 6:  $b = -0.006$ ;  $SE = 0.00$ ;  $p < .05$ ). The spatially lagged version of insider presence is also related to property crimes: 5% fewer property crimes for a one standard deviation increase in the insider presence of contiguous block groups (model 6:  $b = -0.019$ ;  $SE = 0.01$ ;  $p < .01$ ). It therefore appears that the presence of insiders or outsiders is more important to understanding differences in neighborhood crime than the volume of tweets produced by insiders or outsiders. These findings underscore the importance of distinguishing the users of space among traditional and novel measures of the ambient population.

## Discussion

Previous studies have linked administrative measures to crime across neighborhoods. Yet, administrative measures are unable to fully capture the density and type of persons occupying neighborhoods for some period of time. Accordingly, in the present study, we examined spatial predictors of crime derived from Twitter in order to model a dynamic social ecology of crime—most importantly, the composition of the ambient population in the neighborhood. Consistent with studies that reveal differences in collective efficacy, informal social control, and guardianship tendencies (Reynald, 2010, 2011; Wickes et al., 2013), we analytically distinguished between “insiders” and “outsiders,” with the former theorized to have a proclivity for behavioral

intervention whereas the latter more likely to represent uninterested bystanders. We highlight three key findings from the current study.

First, our results showed that the general density of Twitter users in block groups was related to higher counts of violent and property crime, which is consistent with previous research on the ambient population and crime (e.g., Andresen, 2006; Felson & Boivin, 2015). Our results indicate that our Twitter-derived measures of the ambient population may approximate the dynamic nature of human mobility and be useful metrics for the local population at risk (see also Hipp et al., 2019). Indeed, it seems that the crime-enhancing effects of user density is consistent with theoretical arguments that the number of people in an area varies directly with the volume of criminal opportunities (e.g., Cohen & Felson, 1979). That is, a neighborhood with a large ambient population may not only have an increased supply of motivated offenders and potential targets, but a large population may also provide anonymity making it difficult for guardians to detect suspicious behavior (even if there is an increased supply of guardians as well). Relatedly, a higher volume of people generally impedes the development of social ties between residents, store owners, employees, community leaders and other nonresidents who regularly traverse the neighborhood. It therefore can become challenging to establish a shared responsibility for crime and disorder problems in an area. In essence, a larger ambient population is associated with more crime because it is generally believed to offer an abundance of motivated offenders and weakly guarded targets, and to disrupt mechanisms of social organization and control.

Second, the findings demonstrate the significance of distilling the ambient population in ways that are anticipated by theories of informal social control and crime opportunity: indeed, we found that greater insider presence was associated with lower violent and property crime counts. Our insider designation is based on the block group in which a user has tweeted most frequently. As a result, we maintain that any salutary effects of insider presence suggests that people with a conceivably strong vested interest in an area can work together to solve and prevent crime, as well as offset any criminal opportunities that might emerge from the ambient population. While we suggest several mechanisms through which insider presence may be associated with crime, including heightened collective responsibility (e.g., Felson, 1995) and greater effectiveness of guardianship and surveillance (Jacobs, 1961; Reynald & Elffers, 2009), future research might develop methods for further distinguishing between these mechanisms.

More generally, these results indicate that while the ambient population may be important for capturing social ecological processes that impede efforts of informal regulation, accounting for the volume of people in a

neighborhood over some period of time may not be sufficient. Indeed, as Jacobs (1961) warned by drawing on an example of a crowded movie theater (which rarely produces crime): the density of people should not be equated with crime problems. Whether the presence of more people augments or impedes efforts to minimize local crime problems may depend on the composition of people utilizing neighborhood spaces. Accounting for the *composition* of the ambient population in tandem with the size of the ambient population may provide a more comprehensive understanding of neighborhood crime. Future studies might also focus on the composition and diversity of locations outsiders travel from or, in other words, insiders visit. Studying the flows of individuals across spaces might help to shed light on the effects of human mobility on patterns of crime.

Third, our results revealed that the composition of users, whether more insiders or outsiders, matters more for local crime rates than the activity (tweets) of those users. We found that neither tweet density nor insider activity (proportion of tweets made by insiders) were significant predictors of violent or property crime. This may be expected given that measures of tweet activity are likely to capture, at least to some degree, a user contribution bias (Nielsen, 2006). That is, a small proportion of users can account for a large proportion of all tweet activity. Our insider presence metric addresses the user contribution bias by counting each user once in each location (block group) to normalize the effect of the variation in user activities. Nevertheless, the combined results indicate that it is the *presence* of insiders and outsiders that is associated with crime and not necessarily the *activity* of insiders and outsiders.

### *Limitations*

Our study is not without limitations that temper the strength of the findings but also provide potential directions for future research. Specifically, in addition to the limitations discussed at the outset (e.g., sociodemographic representativeness and user contribution bias), there are challenges to our Twitter-derived measures that should be further explored. For instance, while we assume that Twitter users are “insiders” in the areas that they produce the majority of their tweets, it could be that individuals tweet more frequently when they are in places that they are unfamiliar (e.g., concerts or events). In other words, similar to a user contribution bias, there may be a place contribution bias that artificially inflates our estimate of a user’s familiarity in some spatial units and underestimates the user’s familiarity in other spatial units (e.g., residences). Of consequence, it could be that some users are designated as “outsiders” when tweeting from their own home. Prior work,

however, has demonstrated the validity of inferring a user's residence from the location a user tweets most frequently during night-time intervals (Lin & Cromley, 2018); thus, we are confident in our assumption that a user is generally familiar with the location in which they tweet most frequently. Future research might consider estimating our "insider-outsider" distinction in tandem with a "resident" distinction to further decompose the ambient population. Alternative avenues for future studies could be to apply the insider-outsider distinction to ambient populations at specific time-intervals (see Hipp et al., 2019) or to create a measurement procedure that allows for users to be "insiders" of multiple spatial units. Finally, the ability to disaggregate crime incidents by the type of person who committed the crime (insider or outsider) is beyond the scope of the present study. Disaggregating crime into "insider crime" and "outsider crime" may provide additional insight into the mechanisms through which the size and composition of the ambient population is associated with crime (see Boivin & Felson, 2018). For it may be that insiders are more apt to control crime committed by outsiders rather than by those perceived to be insiders of the neighborhood. Such a pattern is consistent with observations of the defended neighborhood which were evident in Suttles' (1968) classic ethnographic study of territoriality dynamics in a near-west side Chicago neighborhood.

## **Conclusion**

Building upon prior work, our study demonstrates how Twitter may be a useful tool for criminologists to measure the ambient population when studying the distribution of crime across neighborhoods (see also Hipp et al., 2019). Further, while prior research has found that crime is generally higher in neighborhoods with larger ambient populations, our current results demonstrate the potential substantive value of partitioning the ambient population into groups with theoretically different interests in the community. Strands of criminological theory suggest that residents and frequent visitors (or "insiders") will have greater responsibility for and investment in the safety and well-being of a community than non-frequent visitors (or "outsiders"). Indeed, consistent with these theoretical assumptions, crime is lower in places where a greater proportion of the ambient population is constituted by insiders—people who presumably have a relatively stronger investment in the well-being of the neighborhood.

## **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors received financial support for the research from the Obermann Center for Advanced Studies at the University of Iowa and the Public Policy Center at the University of Iowa.

## ORCID iD

James C. Wo  <https://orcid.org/0000-0001-9410-0712>

## Notes

1. Stata's likelihood-ratio test, which tests the null hypothesis that the dispersion parameter (alpha) is equal to zero, was significant for all models ( $p < .05$ ).
2. Additional information on the Twitter data and methodology is provided in a supplementary file, available upon request from the corresponding author.

## References

- Achimore, A. (1993). Putting the community back into community retail. *Urban Land*, 52, 33–38.
- Andresen, M. A. (2006). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology*, 46(2), 258–285.
- Andresen, M. A. (2011). The ambient population and crime analysis. *The Professional Geographer*, 63(2), 193–212.
- Andresen, M. A., & Jenion, G. W. (2010). Ambient populations and the calculation of crime rates and risk. *Security Journal*, 23(2), 114–133.
- Bernasco, W., & Block, R. (2011). Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors, and offender anchor points. *Journal of Research in Crime and Delinquency*, 48(1), 33–57.
- Boessen, A., & Hipp, J. R. (2015). Close-ups and the scale of ecology: Land uses and the geography of social context and crime. *Criminology*, 53(3), 399–426.
- Boggs, S. L. (1965). Urban crime patterns. *American Sociological Review*, 30(6), 899–908.
- Boivin, R. (2013). On the use of crime rates. *Canadian Journal of Criminology and Criminal Justice* 55(2), 263–277.
- Boivin, R., & Felson, M. (2018). Crimes by visitors versus crimes by residents: The influence of visitor inflows. *Journal of Quantitative Criminology*, 34(4), 465–480.
- Brantingham, P. J., & Brantingham, P. L. (1984). *Patterns in crime*. MacMillan.
- Brantingham, P. L., & Brantingham, P. J. (1995). Criminality of place. *European Journal on Criminal Policy and Research*, 3: 5–26.
- Browning, C. R., Byron, R. A., Calder, C. A., Krivo, L. J., Kwan, M.-P., Lee, J.-Y., & Peterson, R. D. (2010). Commercial density, residential concentration, and crime: Land use patterns and violence in neighborhood context. *Journal of Research in Crime and Delinquency*, 47(3), 329–357.

- Browning, C. R., Calder, C. A., Boettner, B., & Smith, A. (2017). Ecological networks and urban crime: The structure of shared routine activity locations and neighborhood-level informal control capacity. *Criminology*, 55(4), 754–778.
- Bursik, R. J. (1999). The informal control of crime through neighborhood networks. *Sociological Focus*, 32(1), 85–97.
- Bursik, R. J., & Grasmick, H. G. (1993). *Neighborhoods and crime: The dimensions of effective community control*. Lexington Books.
- Chamlin, M. B., & Cochran, J. K. (2004). An excursus on the population size-crime relationship. *Western Criminology Review*, 5(2), 119–130.
- Clarke, R. V. (1992). *Situational crime prevention: Successful case studies*. Harrow and Heston.
- Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 44(4), 588–608.
- Cohen, L. E., Kaufman, R. L., & Gottfredson, M. R. (1985). Risk-based crime statistics: A forecasting comparison for burglary and auto theft. *Journal of Criminal Justice*, 13(5), 445–457.
- Felson, M. (1987). Routine activities and crime prevention in the developing metropolis\*. *Criminology*, 25(4), 911–932.
- Felson, M. (1995). Those who discourage crime. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 53–66). Willow Tree Press.
- Felson, M., & Boba, R. (2010). *Crime and everyday life* (4th ed.). SAGE.
- Felson, M., & Boivin, R. (2015). Daily crime flows within a city. *Crime Science*, 4(1), 31.
- Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*, 61(1), 115–125.
- Gibbs, J. P., & Martin, W. T. (1962). Urbanization, technology, and the division of labor: International patterns. *American Sociological Review*, 27(5), 667–677.
- Greenberg, S. W., Rohe, W. M., & Williams, J. R. (1982). Safety in urban neighborhoods: A comparison of physical characteristics and informal territorial control in high and low crime neighborhoods. *Population and Environment*, 5, 141–165.
- Haberman, C. P., & Ratcliffe, J. H. (2015). Testing for temporally differentiated relationships among potentially criminogenic places and census block street robbery counts. *Criminology*, 53(3), 457–483.
- Hanaoka, K. (2018). New insights on relationships between street crimes and ambient population: Use of hourly population data estimated from mobile phone users' locations. *Environment and Planning B: Urban Analytics and City Science*, 45(2), 295–311.
- Hilbe, J. M. (2007). *Negative binomial regression*. Cambridge University Press.
- Hillier, B. (2004). Can streets be made safe? *Urban Design International*, 9(1), 31–45.
- Hipp, J. R., Bates, C., Lichman, M., & Smyth, P. (2019). Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly*, 36(4), 718–748.



- Hipp, J. R., & Roussell, A. (2013). Micro- and macro-environment population and the consequences for crime rates. *Social Forces*, 92(2), 563–595.
- Hipp, J. R., & Williams, S. A. (2020). Advances in spatial criminology: The spatial scale of crime. *Annual Review of Criminology*, 3(1), 75–95.
- Hipp, J. R., Wo, J. C., & Kim, Y.-A. (2017). Studying neighborhood crime across different macro spatial scales: The case of robbery in 4 cities. *Social Science Research*, 68, 15–29.
- Hollis-Peel, M. E., Reynald, D. M., van Bavel, M., Elffers, H., & Welsh, B. C. (2011). Guardianship for crime prevention: A critical review of the literature. *Crime, Law and Social Change*, 56, 53–70.
- Hunter, A. (1985). Private, parochial and public social orders: The problem of crime and incivility in urban communities. In G. D. Suttles & M. N. Zald (Eds.), *The challenge of social control: Citizenship and institution building in modern society* (pp. 230–242). Ablex.
- Jacobs, J. (1961). *The death and life of great American cities*. Vintage Books.
- Kubrin, C. E., & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime & Delinquency*, 40(4), 374–402.
- Leetaru, K. (2019). *The era of precision mapping of social media is coming to an end*. Forbes. <https://www.forbes.com/sites/kalevleetaru/2019/03/06/the-era-of-precision-mapping-of-social-media-is-coming-to-an-end/?sh=64f9ad715824>
- Lin, J., & Cromley, R. G. (2018). Inferring the home locations of Twitter users based on the spatiotemporal clustering of Twitter data. *Transactions in GIS*, 22(1), 82–97.
- Malleson, N., & Andresen, M. A. (2015a). The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, 42(2), 112–121.
- Malleson, N., & Andresen, M. A. (2015b). Spatio-temporal crime hotspots and the ambient population. *Crime Science*, 4(1), 10.
- Malleson, N., & Andresen, M. A. (2016). Exploring the impact of ambient population measures on London crime hotspots. *Journal of Criminal Justice*, 46, 52–63.
- Mburu, L. W., & Helbich, M. (2016). Crime risk estimation with a commuter-harmonized ambient population. *Annals of the American Association of Geographers*, 106(4), 804–818.
- McCord, E. S., Ratcliffe, J. H., Garcia, R. M., & Taylor, R. B. (2007). Nonresidential crime attractors and generators elevate perceived neighborhood crime and incivilities. *Journal of Research in Crime and Delinquency*, 44(3), 295–320.
- Newman, O. (1972). *Defensible space: Crime prevention through urban design*. Macmillan.
- Nielsen, J. (2006). *Participation inequality: Encouraging more users to contribute*. Nielsen Norman Group. <https://www.nngroup.com/articles/participation-inequality>
- Oldenburg, R. (1999). *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community*. Marlowe & Company.

- Osgood, D. W., & Chambers, J. M. (2000). Social disorganization outside the Metropolis: An Analysis of rural youth violence\*. *Criminology*, 38(1), 81–116.
- Pavalanathan, U., & Eisenstein, J. (2015). Confounds and consequences in geotagged twitter data. *arXiv preprint arXiv:1506.02275*.
- Peterson, R. D., & Krivo, L. J. (2010). *Divergent social worlds: Neighborhood crime and the racial-spatial divide*. Russell Sage Foundation.
- Pratt, T. C., & Cullen, F. T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. In M. Tonry (Ed.), *Crime and Justice: A Review of Research* (pp. 373–450). The University of Chicago Press.
- Quattrone, G., Capra, L., & De Meo, P. (2015). There's no such thing as the perfect map: Quantifying bias in spatial crowd-sourcing datasets. In D. Cosley, A. Forte, L. Ciolfi, & D. McDonalld (Eds.), *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1021–1032). ACM, Vancouver, Canada.
- Reynald, D. M. (2010). Guardians on guardianship: Factors affecting the willingness to supervise, the ability to detect potential offenders, and the willingness to intervene. *Journal of Research in Crime and Delinquency*, 47(3), 358–390.
- Reynald, D. M. (2011). Factors associated with the guardianship of places: Assessing the relative importance of the spatio-physical and sociodemographic contexts in generating opportunities for capable guardianship. *Journal of Research in Crime and Delinquency*, 48(1), 110–142.
- Reynald, D. M., & Elffers, H. (2009). The future of Newman's defensible space theory: Linking defensible space and the routine activities of place. *European Journal of Criminology*, 6(1), 25–46.
- Roncek, D. W. (1981). Dangerous places: Crime and residential environment. *Social Forces*, 60(1), 74–96.
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94(4), 774–802.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918–924.
- Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. University of Chicago Press.
- Song, G., Liu, L., Bernasco, W., Xiao, L., Zhou, S., & Liao, W. (2018). Testing indicators of risk populations for theft from the person across space and time: The significance of mobility and outdoor activity. *Annals of the American Association of Geographers*, 108(5), 1370–1388.
- Stucky, T. D., & Ottensmann, J. R. (2009). Land use and violent crime. *Criminology*, 47(7), 1223–1264.
- Stults, B. J., & Hasbrouck, M. (2015). The effect of commuting on city-level crime rates. *Journal of Quantitative Criminology*, 31(2), 331–350.
- Suttles, G. D. (1968). *The social order of the slum: Ethnicity and territory in the inner city*. University of Chicago Press.
- Talen, E. (2002). The social goals of new urbanism. *Housing Policy Debate*, 13(1), 165–188.

- Taylor, R. B., Koons, B. A., Kurtz, E. M., Greene, J. R., & Perkins, D. D. (1995). Street blocks with more nonresidential land use have more physical deterioration: Evidence from Baltimore and Philadelphia. *Urban Affairs Review*, 31(1), 120–136.
- Tsou, M.-H. (2015). Research challenges and opportunities in mapping social media and big data. *Cartography and Geographic Information Science*, 42(1), 70–74.
- Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *ICWSM*, 14, 505–514.
- Wang, M., & Gerber, M. S. (2015). *Using Twitter for next-place prediction, with an application to crime prediction*. 2015 IEEE Symposium Series on Computational Intelligence [Conference session], Cape Town, South Africa (pp. 941–948).
- Wang, X., Brown, D. E., & Gerber, M. S. (2012). *Spatio-temporal modeling of criminal incidents using geographic, demographic, and twitter-derived information*. 2012 IEEE International Conference on Intelligence and Security Informatics, Washington, D.C., USA, (pp. 36–41).
- Wickes, R., Hipp, J. R., Sargeant, E., & Homel, R. (2013). Collective efficacy as a task specific process: Examining the relationship between social ties, neighborhood cohesion and the capacity to Respond to violence, delinquency and civic problems. *American Journal of Community Psychology*, 52(1–2), 115–127.
- Wilcox, P., Quisenberry, N., Cabrera, D. T., & Jones, S. (2004). Busy places and broken windows? Toward defining the role of physical structure and process in community crime models. *Sociological Quarterly*, 45(2), 185–207.
- Wo, J. C. (2016). Community context of crime: A longitudinal examination of the effects of local institutions on neighborhood crime. *Crime & Delinquency*, 62(10), 1286–1312.
- Wo, J. C. (2019a). Mixed land use and neighborhood crime. *Social Science Research*, 78, 170–186.
- Wo, J. C. (2019b). Understanding the differential effects of land uses on crime: An examination across Philadelphia neighborhoods. *British Journal of Criminology*, 59(6), 1432–1454.
- Zahnow, R. (2018). Mixed land use: Implications for violence and property crime. *City & Community*, 17(4), 1119–1142.

## Author Biographies

**James C. Wo** is an assistant professor in the Department of Sociology and Criminology at the University of Iowa, and a research fellow at the University of Iowa Public Policy Center. His research interests include neighborhoods and crime, local institutions and organizations, land uses, and quantitative research methods.

**Ethan M. Rogers** is a research associate in the Crime and Justice Policy Research Program at the University of Iowa Public Policy Center. His research interests include conflict escalation and interpersonal violence, the mobilization of the law, the neighborhood context of crime, and criminal justice processing.

**Mark T. Berg** is an associate professor in the Department of Sociology and Criminology and director of the Crime and Justice Policy Research Program at the Public Policy Center at the University of Iowa. His current research interests include the social psychology of interpersonal violence and the social contexts of health and well-being across the life course.

**Caglar Koynu** is an assistant professor in the Department of Geographical and Sustainability Sciences at the University of Iowa, and a research fellow at the University of Iowa Public Policy Center. His research interests include analysis and visualization of large geo-social networks, that is, networks embedded in geographic space and time such as migration, human mobility, commodity flows, and information flows from social media.